

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2002, 18:03:43 ; Search time 20.94 Seconds
(without alignments)
1670.321 Million cell updates/sec

Title: US-09-886-400-4

Perfect score: 1877

Sequence: 1 LRALVFHGNLQYAEIPKSEI.....RLDAFRAIYDMGENCEP 364

Scoring table: BLOSUM62

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1494	79.6	364	2	E71144
2	169	9.0	529	2	AC2112
3	162.5	8.7	529	2	S76831
4	161.5	8.6	633	2	G71241
5	157.5	8.4	686	1	ALDYAT
6	147	7.8	527	2	E71297
7	145.5	7.8	655	2	E75206
8	141.5	7.5	467	2	B64501
9	141.5	7.5	649	2	A49512
10	130.5	7.0	560	2	D71011
11	128	6.8	923	2	G83826
12	124	6.6	1362	2	A75207
13	123.5	6.6	744	2	AC1970
14	120.5	6.4	324	1	B69553
15	119	6.3	602	2	C75120
16	113	6.0	1069	2	AF1930
17	112	6.0	312	2	I40383
18	111	5.9	426	2	D71334
19	109.5	5.8	447	2	E90250
20	106	5.6	526	2	B70859
21	105.5	5.6	902	2	E90270
22	104	5.5	704	2	G96587
23	103.5	5.5	314	2	F84044
24	101.5	5.4	705	2	F70475
25	100.5	5.4	266	2	G96943
26	100.5	5.4	619	2	D71361
27	99	5.3	4845	2	T31067
28	98.5	5.2	588	2	AE1926
29	98	5.2	503	2	T19319

ALIGNMENTS

30	98	5.2	764	1	H70414	conserved hypotet
31	97.5	5.2	402	2	T14710	probable transpos
32	97.5	5.2	402	2	AB0193	transposase, 15285
33	97.5	5.2	402	2	AB0358	transposase, 15285
34	97.5	5.2	402	2	AB0472	transposase, 15285
35	97.5	5.2	402	2	AC0341	transposase, 15285
36	97.5	5.2	402	2	AD0002	transposase, 15285
37	97.5	5.2	402	2	AE0190	transposase, 15285
38	97.5	5.2	402	2	AE0242	transposase, 15285
39	97.5	5.2	402	2	AE0267	transposase, 15285
40	97.5	5.2	402	2	AF0348	transposase, 15285
41	97.5	5.2	402	2	AG0210	transposase, 15285
42	97.5	5.2	402	2	AG0260	transposase, 15285
43	97.5	5.2	402	2	AG0264	transposase, 15285
44	97.5	5.2	402	2	AG0339	transposase, 15285
45	97.5	5.2	402	2	AH0193	transposase, 15285

RESULT 1

E71144
hypothetical protein PH0368 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 08-Sep-2000

C:Accession: E71144

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Ogi

M.; Onofuku, Y.; Funahashi, T.; Tanaka, T.; Kidoh, Y.; Yamazaki, J.; Kushida, N.; Ogi

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic

A:Reference number: A71000; MUID:98344137

A:Accession: E71144

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-364 <RAW>

A:Cross-references: GB:AP000002; NID:g3236129; PIDN:BA29442.1; PID:d1030385; PID:g32

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by Genba

C:Genetics:

C:Gene: PH0368

C:Superfamily: Pyrococcus horikoshii hypothetical protein PH0368

Query Match	79.6%	Score 1494;	DB 2;	Length 364;
Best Local Similarity	75.8%	Pred. No. 2.7e-105;		
Matches 275;	Conservative 49;	Mismatches 39;	Indels 0;	Gaps 0;
1	LRALVFHGNLQYAEIPKSEI	PKYIEKAYIPVETLIKKEIIPGINTGYTKLPDID	60	
1	MRALIFHGNLQYAEIPKSEI	PKYIEKAYIPVETLIKKEIIPGINTGYTKLPDID	60	
61	LVKGGISDLIEITIGSYTAI	PLPLSRVEAOVORDREKELPELSKGGWLELAY	120	
61	LVKGGISDLIEITIGSYTAI	PLPLSRVEAOVORDREKELPELSKGGWLELAY	120	
61	LVKGGISDLIEITIGSYTAI	PLPLSRVEAOVORDREKELPELSKGGWLELAY	120	
121	DPILPATIKDNGYETLPADE	GFAMLSAHSAPKIPKLPPLIKAKQERFRYSITLG	180	
121	DPILPATIKDNGYETLPADE	GFAMLSAHSAPKIPKLPPLIKAKQERFRYSITLG	180	
121	DPILPATIKDNGYETLPADE	GFAMLSAHSAPKIPKLPPLIKAKQERFRYSITLG	180	
121	DPILPATIKDNGYETLPADE	GFAMLSAHSAPKIPKLPPLIKAKQERFRYSITLG	180	
181	LRLEKRAIKVEEGKVT	LKAVKDIKAVVAVNTAVMLIGRLPLMPKVASWIEDK	240	
181	LRLEKRAIKVEEGKVT	LKAVKDIKAVVAVNTAVMLIGRLPLMPKVASWIEDK	240	
181	LRLEKRAIKVEEGKVT	LKAVKDIKAVVAVNTAVMLIGRLPLMPKVASWIEDK	240	
181	LRLEKRAIKVEEGKVT	LKAVKDIKAVVAVNTAVMLIGRLPLMPKVASWIEDK	240	
241	NILKGYDIEITIGSYTAI	PLPLSRVEAOVORDREKELPELSKGGWLELAY	120	
241	NILKGYDIEITIGSYTAI	PLPLSRVEAOVORDREKELPELSKGGWLELAY	120	
241	NILKGYDIEITIGSYTAI	PLPLSRVEAOVORDREKELPELSKGGWLELAY	120	
241	NILKGYDIEITIGSYTAI	PLPLSRVEAOVORDREKELPELSKGGWLELAY	120	
301	DKSLIRVEDEGNARLN	LSYNGEGLLEALLAENS	DARGWEPLPRRLDAFRAIYDMRGE	360
301	DKSLIRVEDEGNARLN	LSYNGEGLLEALLAENS	DARGWEPLPRRLDAFRAIYDMRGE	360
301	DKSLIRVEDEGNARLN	LSYNGEGLLEALLAENS	DARGWEPLPRRLDAFRAIYDMRGE	360
301	DKSLIRVEDEGNARLN	LSYNGEGLLEALLAENS	DARGWEPLPRRLDAFRAIYDMRGE	360
361	NCE 363			

Db 361 NGK 363

RESULT 2

AC2112 hypothetical protein alr2450 (imported) - Anabaena sp. (strain PCC 7120)

C:Species: Anabaena sp.

C:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002

C:Accession: AC2112

R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iriuch

Nakazaki, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AC2112

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-529 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAW74149.1; PID:g17131542; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:gene: alr2450

C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1386

Query Match 9.0%; Score 169; DB 2; Length 529;

Best Local Similarity 21.0%; Pred. No. 5.6e-05;

Matches 104; Conservative 55; Mismatches 162; Indels 174; Gaps 21;

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C:Accession: S76831

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,

O., K.; Okumura, S.; Shimizu, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys

S.

A:Reference number: S74322; MUID:97061201

A:Accession: S76831

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-529 <KAN>

A:Cross-references: EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BA18743.1; PID:g165

A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1386

Query Match 8.7%; Score 162.5; DB 2; Length 529;

Best Local Similarity 21.4%; Pred. No. 0.00017;

Matches 101; Conservative 47; Mismatches 140; Indels 183; Gaps 19;

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

Db 361 NGK 363

```

QY 1 LRAIVFNGNLQYAEIPKSEIPKVIKAVIPVETLKEEIPFGINTGYTLKFLPKDII 60
Db 1 LRAIVFNGNLQYAEIPKSEIPKVIKAVIPVETLKEEIPFGINTGYTLKFLPKDII 60
QY 61 LVKGASDLLEIIGTSTHAILPLPSRVEAQVQDRREVEKEELFELSPKGFMLPELAY 120
Db 61 LVKGASDLLEIIGTSTHAILPLPSRVEAQVQDRREVEKEELFELSPKGFMLPELAY 120
QY 121 DPTIPALIKNDNGEYELFADGEMLFSAHLSAIPKIPKIPVPHILKAKREKFRISTYLLG 180
Db 121 DPTIPALIKNDNGEYELFADGEMLFSAHLSAIPKIPKIPVPHILKAKREKFRISTYLLG 180
QY 181 LREIRKAIKLVFEGKVTLLKAVKDIEAVPVWVAVNTAVMLGIGRLPLNPKKVASWIEDKD 240
Db 181 LREIRKAIKLVFEGKVTLLKAVKDIEAVPVWVAVNTAVMLGIGRLPLNPKKVASWIEDKD 240
QY 241 NILVGTDIIEFTGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
Db 241 NILVGTDIIEFTGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
QY 301 DKSIRIRREDEGNARLNLSTYMRGELALLAENSADARGWEPLPERRLDAFRATYNDWGE 360
Db 301 DKSIRIRREDEGNARLNLSTYMRGELALLAENSADARGWEPLPERRLDAFRATYNDWGE 360
QY 361 NCEP 364
Db 361 NCEP 364

```

```

RESULT 2
US-09-407-806-4
; Sequence 4, Application US/09407806
; GENERAL INFORMATION:
; APPLICANT: Murphy, Dennis
; TITLE OF INVENTION: ALPHA-GALACTOSIDASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/407,806
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/613,220
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5070
; TELEFAX: 619-68-5099
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 346 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
US-09-407-806-4

```

```

Query Match 84.2%; Score 1580; DB 18; Length 346;
Best Local Similarity 94.5%; Pred. No. 3,4e-152;
Matches 344; Conservative 1; Mismatches 1; Indels 18; Gaps 18;
QY 1 LRAIVFNGNLQYAEIPKSEIPKVIKAVIPVETLKEEIPFGINTGYTLKFLPKDII 60
Db 1 LRAIVFNGNLQYAEIPKSEIPKVIKAVIPVETLKEEIPFGINTGYTLKFLPKDII 57
QY 61 LVKGASDLLEIIGTSTHAILPLPSRVEAQVQDRREVEKEELFELSPKGFMLPELAY 120
Db 58 LVKGASDLLEIIGTSTHAILPLPSRVEAQVQDRREVEKEELFELSPKGFMLPELAY 114
QY 121 DPTIPALIKNDNGEYELFADGEMLFSAHLSAIPKIPKIPVPHILKAKREKFRISTYLLG 180
Db 115 DPTIPALIKNDNGEYELFADGEMLFSAHLSAIPKIPKIPVPHILKAKREKFRISTYLLG 171
QY 181 LREIRKAIKLVFEGKVTLLKAVKDIEAVPVWVAVNTAVMLGIGRLPLNPKKVASWIEDKD 240
Db 172 LREIRKAIKLVFEGKVTLLKAVKDIEAVPVWVAVNTAVMLGIGRLPLNPKKVASWIEDKD 228
QY 241 NILVGTDIIEFTGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
Db 229 NILVGTDIIEFTGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 285
QY 301 DKSIRIRREDEGNARLNLSTYMRGELALLAENSADARGWEPLPERRLDAFRATYNDWGE 360
Db 286 DKSIRIRREDEGNARLNLSTYMRGELALLAENSADARGWEPLPERRLDAFRATYNDWGE 342
QY 361 NCEP 364
Db 343 NCEP 346

```

```

RESULT 3
US-09-407-806A-4
; Sequence 4, Application US/09407806A
; GENERAL INFORMATION:
; APPLICANT: Murphy, Dennis
; TITLE OF INVENTION: ALPHA-GALACTOSIDASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/407,806A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/613,220
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5070
; TELEFAX: 619-68-5099
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 346 amino acids
; TYPE: amino acid

```

-

